



IFWO

RAW SEQUENCE LISTING

DATE: 08/19/2004

PATENT APPLICATION: US/10/801,266

TIME: 09:54:38

Input Set : A:\PTO.LM.txt

Output Set : N:\CRF4\08192004\J801266.raw

3 <110> APPLICANT: Structural GenomiX, Inc.
 4 Atwell, Shane
 5 Buchanan, Sean G.
 7 <120> TITLE OF INVENTION: Crystals and Structures of Epidermal Growth Factor Receptor
 Kinase Domain
 9 <130> FILE REFERENCE: 022132-000910US
 11 <140> CURRENT APPLICATION NUMBER: 10/801,266
 12 <141> CURRENT FILING DATE: 2004-03-16
 14 <150> PRIOR APPLICATION NUMBER: US 60/456,553
 15 <151> PRIOR FILING DATE: 2003-03-20
 17 <160> NUMBER OF SEQ ID NOS: 8
 19 <170> SOFTWARE: PatentIn version 3.1
 21 <210> SEQ ID NO: 1
 22 <211> LENGTH: 18
 23 <212> TYPE: DNA
 24 <213> ORGANISM: Artificial Sequence
 26 <220> FEATURE:
 27 <223> OTHER INFORMATION: forward primer
 29 <400> SEQUENCE: 1
 30 gctcccaacc aagctctc 18
 33 <210> SEQ ID NO: 2
 34 <211> LENGTH: 20
 35 <212> TYPE: DNA
 36 <213> ORGANISM: Artificial Sequence
 38 <220> FEATURE:
 39 <223> OTHER INFORMATION: reverse primer
 41 <400> SEQUENCE: 2
 42 cccctggaat gacaaggtag 20
 45 <210> SEQ ID NO: 3
 46 <211> LENGTH: 26
 47 <212> TYPE: DNA
 48 <213> ORGANISM: Artificial Sequence
 50 <220> FEATURE:
 51 <223> OTHER INFORMATION: sequence after being ligated into vector
 53 <400> SEQUENCE: 3
 54 aagggcacatca tcaccatcac cactga 26
 57 <210> SEQ ID NO: 4
 58 <211> LENGTH: 297
 59 <212> TYPE: PRT
 60 <213> ORGANISM: Artificial Sequence
 62 <220> FEATURE:
 63 <223> OTHER INFORMATION: predicted sequence of EGFRKD expressed protein
 65 <400> SEQUENCE: 4
 67 Met Ala Leu Ala Pro Asn Gln Ala Leu Leu Arg Ile Leu Lys Glu Thr



RAW SEQUENCE LISTING

DATE: 08/19/2004

PATENT APPLICATION: US/10/801,266

TIME: 09:54:38

Input Set : A:\PTO.LM.txt

Output Set : N:\CRF4\08192004\J801266.raw

```

68 1          5          10          15
71 Glu Phe Lys Lys Ile Lys Val Leu Gly Ser Gly Ala Phe Gly Thr Val
72          20          25          30
75 Tyr Lys Gly Leu Trp Ile Pro Glu Gly Glu Lys Val Lys Ile Pro Val
76          35          40          45
79 Ala Ile Lys Glu Leu Arg Glu Ala Thr Ser Pro Lys Ala Asn Lys Glu
80          50          55          60
83 Ile Leu Asp Glu Ala Tyr Val Met Ala Ser Val Asp Asn Pro His Val
84 65          70          75          80
87 Cys Arg Leu Leu Gly Ile Cys Leu Thr Ser Thr Val Gln Leu Ile Thr
88          85          90          95
91 Gln Leu Met Pro Phe Gly Cys Leu Leu Asp Tyr Val Arg Glu His Lys
92          100          105          110
95 Asp Asn Ile Gly Ser Gln Tyr Leu Leu Asn Trp Cys Val Gln Ile Ala
96          115          120          125
99 Lys Gly Met Asn Tyr Leu Glu Asp Arg Arg Leu Val His Arg Asp Leu
100          130          135          140
103 Ala Ala Arg Asn Val Leu Val Lys Thr Pro Gln His Val Lys Ile Thr
104 145          150          155          160
107 Asp Phe Gly Leu Ala Lys Leu Leu Gly Ala Glu Glu Lys Glu Tyr His
108          165          170          175
111 Ala Glu Gly Gly Lys Val Pro Ile Lys Trp Met Ala Leu Glu Ser Ile
112          180          185          190
115 Leu His Arg Ile Tyr Thr His Gln Ser Asp Val Trp Ser Tyr Gly Val
116          195          200          205
119 Thr Val Trp Glu Leu Met Thr Phe Gly Ser Lys Pro Tyr Asp Gly Ile
120          210          215          220
123 Pro Ala Ser Glu Ile Ser Ser Ile Leu Glu Lys Gly Glu Arg Leu Pro
124 225          230          235          240
127 Gln Pro Pro Ile Cys Thr Ile Asp Val Tyr Met Ile Met Val Lys Cys
128          245          250          255
131 Trp Met Ile Asp Ala Asp Ser Arg Pro Lys Phe Arg Glu Leu Ile Ile
132          260          265          270
135 Glu Phe Ser Lys Met Ala Arg Asp Pro Gln Arg Tyr Leu Val Ile Gln
136          275          280          285
139 Gly Glu Gly His His His His His His
140          290          295
143 <210> SEQ ID NO: 5
144 <211> LENGTH: 290
145 <212> TYPE: PRT
146 <213> ORGANISM: Homo sapiens
148 <400> SEQUENCE: 5
150 Met Ala Leu Ala Pro Asn Gln Ala Leu Leu Arg Ile Leu Lys Glu Thr
151 1          5          10          15
154 Glu Phe Lys Lys Ile Lys Val Leu Gly Ser Gly Ala Phe Gly Thr Val
155          20          25          30
158 Tyr Lys Gly Leu Trp Ile Pro Glu Gly Glu Lys Val Lys Ile Pro Val
159          35          40          45
162 Ala Ile Lys Glu Leu Arg Glu Ala Thr Ser Pro Lys Ala Asn Lys Glu

```

RAW SEQUENCE LISTING

DATE: 08/19/2004

PATENT APPLICATION: US/10/801,266

TIME: 09:54:38

Input Set : A:\PTO.LM.txt

Output Set: N:\CRF4\08192004\J801266.raw

```

163      50      55      60
166 Ile Leu Asp Glu Ala Tyr Val Met Ala Ser Val Asp Asn Pro His Val
167 65      70      75      80
170 Cys Arg Leu Leu Gly Ile Cys Leu Thr Ser Thr Val Gln Leu Ile Thr
171      85      90      95
174 Gln Leu Met Pro Phe Gly Cys Leu Leu Asp Tyr Val Arg Glu His Lys
175      100      105      110
178 Asp Asn Ile Gly Ser Gln Tyr Leu Leu Asn Trp Cys Val Gln Ile Ala
179      115      120      125
182 Lys Gly Met Asn Tyr Leu Glu Asp Arg Arg Leu Val His Arg Asp Leu
183      130      135      140
186 Ala Ala Arg Asn Val Leu Val Lys Thr Pro Gln His Val Lys Ile Thr
187 145      150      155      160
190 Asp Phe Gly Leu Ala Lys Leu Leu Gly Ala Glu Glu Lys Glu Tyr His
191      165      170      175
194 Ala Glu Gly Gly Lys Val Pro Ile Lys Trp Met Ala Leu Glu Ser Ile
195      180      185      190
198 Leu His Arg Ile Tyr Thr His Gln Ser Asp Val Trp Ser Tyr Gly Val
199      195      200      205
202 Thr Val Trp Glu Leu Met Thr Phe Gly Ser Lys Pro Tyr Asp Gly Ile
203      210      215      220
206 Pro Ala Ser Glu Ile Ser Ser Ile Leu Glu Lys Gly Glu Arg Leu Pro
207 225      230      235      240
210 Gln Pro Pro Ile Cys Thr Ile Asp Val Tyr Met Ile Met Val Lys Cys
211      245      250      255
214 Trp Met Ile Asp Ala Asp Ser Arg Pro Lys Phe Arg Glu Leu Ile Ile
215      260      265      270
218 Glu Phe Ser Lys Met Ala Arg Asp Pro Gln Arg Tyr Leu Val Ile Gln
219      275      280      285
222 Gly Glu
223      290
226 <210> SEQ ID NO: 6
227 <211> LENGTH: 289
228 <212> TYPE: PRT
229 <213> ORGANISM: Unknown
231 <220> FEATURE:
232 <223> OTHER INFORMATION: HER2/ErbB2
234 <400> SEQUENCE: 6
236 Ser Gly Ala Met Pro Asn Gln Ala Gln Met Arg Ile Leu Lys Glu Thr
237 1      5      10      15
240 Glu Leu Arg Lys Val Lys Val Leu Gly Ser Gly Ala Phe Gly Thr Val
241      20      25      30
244 Tyr Lys Gly Ile Trp Ile Pro Asp Gly Glu Asn Val Lys Ile Pro Val
245      35      40      45
248 Ala Ile Lys Val Leu Arg Glu Asn Thr Ser Pro Lys Ala Asn Lys Glu
249      50      55      60
252 Ile Leu Asp Glu Ala Tyr Val Met Ala Gly Val Gly Ser Pro Tyr Val
253 65      70      75      80
256 Ser Arg Leu Leu Gly Ile Cys Leu Thr Ser Thr Val Gln Leu Val Thr

```

RAW SEQUENCE LISTING

DATE: 08/19/2004

PATENT APPLICATION: US/10/801,266

TIME: 09:54:38

Input Set : A:\PTO.LM.txt

Output Set: N:\CRF4\08192004\J801266.raw

```

257          85          90          95
260 Gln Leu Met Pro Tyr Gly Cys Leu Leu Asp His Val Arg Glu Asn Arg
261          100          105          110
264 Gly Arg Leu Gly Ser Gln Asp Leu Leu Asn Trp Cys Met Gln Ile Ala
265          115          120          125
268 Lys Gly Met Ser Tyr Leu Glu Asp Val Arg Leu Val His Arg Asp Leu
269          130          135          140
272 Ala Ala Arg Asn Val Leu Val Lys Ser Pro Asn His Val Lys Ile Thr
273 145          150          155          160
276 Asp Phe Gly Leu Ala Arg Leu Leu Asp Ile Asp Glu Thr Glu Tyr His
277          165          170          175
280 Ala Asp Gly Gly Lys Val Pro Ile Lys Trp Met Ala Leu Glu Ser Ile
281          180          185          190
284 Leu Arg Arg Arg Phe Thr His Gln Ser Asp Val Trp Ser Tyr Gly Val
285          195          200          205
288 Thr Val Trp Glu Leu Met Thr Phe Gly Ala Lys Pro Tyr Asp Gly Ile
289          210          215          220
292 Pro Ala Arg Glu Ile Pro Asp Leu Leu Glu Lys Gly Glu Arg Leu Pro
293 225          230          235          240
296 Gln Pro Pro Ile Cys Thr Ile Asp Val Tyr Met Ile Met Val Lys Cys
297          245          250          255
300 Trp Met Ile Asp Ser Glu Cys Arg Pro Arg Phe Arg Glu Leu Val Ser
301          260          265          270
304 Glu Phe Ser Arg Met Ala Arg Asp Pro Gln Arg Phe Val Val Ile Gln
305          275          280          285
308 Asn
312 <210> SEQ ID NO: 7
313 <211> LENGTH: 290
314 <212> TYPE: PRT
315 <213> ORGANISM: Unknown
317 <220> FEATURE:
318 <223> OTHER INFORMATION: HER4/ErbB4
320 <400> SEQUENCE: 7
322 Ser Gly Thr Ala Pro Asn Gln Ala Gln Leu Arg Ile Leu Lys Glu Thr
323 1          5          10          15
326 Glu Leu Lys Arg Val Lys Val Leu Gly Ser Gly Ala Phe Gly Thr Val
327          20          25          30
330 Tyr Lys Gly Ile Trp Val Pro Glu Gly Glu Thr Val Lys Ile Pro Val
331          35          40          45
334 Ala Ile Lys Ile Leu Asn Glu Thr Thr Gly Pro Lys Ala Asn Val Glu
335          50          55          60
338 Phe Met Asp Glu Ala Leu Ile Met Ala Ser Met Asp His Pro His Leu
339 65          70          75          80
342 Val Arg Leu Leu Gly Val Cys Leu Ser Pro Thr Ile Gln Leu Val Thr
343          85          90          95
346 Gln Leu Met Pro His Gly Cys Leu Leu Glu Tyr Val His Glu His Lys
347          100          105          110
350 Asp Asn Ile Gly Ser Gln Leu Leu Leu Asn Trp Cys Val Gln Ile Ala
351          115          120          125

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/801,266

DATE: 08/19/2004

TIME: 09:54:38

Input Set : A:\PTO.LM.txt

Output Set: N:\CRF4\08192004\J801266.raw

```

354 Lys Gly Met Met Tyr Leu Glu Glu Arg Arg Leu Val His Arg Asp Leu
355      130      135      140
358 Ala Ala Arg Asn Val Leu Val Lys Ser Pro Asn His Val Lys Ile Thr
359 145      150      155      160
362 Asp Phe Gly Leu Ala Arg Leu Leu Glu Gly Asp Glu Lys Glu Tyr Asn
363      165      170      175
366 Ala Asp Gly Gly Lys Met Pro Ile Lys Trp Met Ala Leu Glu Cys Ile
367      180      185      190
370 His Tyr Arg Lys Phe Thr His Gln Ser Asp Val Trp Ser Tyr Gly Val
371      195      200      205
374 Thr Ile Trp Glu Leu Met Thr Phe Gly Gly Lys Pro Tyr Asp Gly Ile
375      210      215      220
378 Pro Thr Arg Glu Ile Pro Asp Leu Leu Glu Lys Gly Glu Arg Leu Pro
379 225      230      235      240
382 Gln Pro Pro Ile Cys Thr Ile Asp Val Tyr Met Val Met Val Lys Cys
383      245      250      255
386 Trp Met Ile Asp Ala Asp Ser Arg Pro Lys Phe Lys Glu Leu Ala Ala
387      260      265      270
390 Glu Phe Ser Arg Met Ala Arg Asp Pro Gln Arg Tyr Leu Val Ile Gln
391      275      280      285
394 Gly Asp
395      290
398 <210> SEQ ID NO: 8
399 <211> LENGTH: 289
400 <212> TYPE: PRT
401 <213> ORGANISM: Unknown
403 <220> FEATURE:
404 <223> OTHER INFORMATION: HER3/ErbB3
406 <400> SEQUENCE: 8
408 Ser Glu Lys Ala Asn Lys Val Leu Ala Arg Ile Phe Lys Glu Thr Glu
409 1      5      10      15
412 Leu Arg Lys Leu Lys Val Leu Gly Ser Gly Val Phe Gly Thr Val His
413      20      25      30
416 Lys Gly Val Trp Ile Pro Glu Gly Glu Ser Ile Lys Ile Pro Val Cys
417      35      40      45
420 Ile Lys Val Ile Glu Asp Lys Ser Gly Arg Gln Ser Phe Gln Ala Val
421      50      55      60
424 Thr Asp His Met Leu Ala Ile Gly Ser Leu Asp His Ala His Ile Val
425 65      70      75      80
428 Arg Leu Leu Gly Leu Cys Pro Gly Ser Ser Leu Gln Leu Val Thr Gln
429      85      90      95
432 Tyr Leu Pro Leu Gly Ser Leu Leu Asp His Val Arg Gln His Arg Gly
433      100      105      110
436 Ala Leu Gly Pro Gln Leu Leu Leu Asn Trp Gly Val Gln Ile Ala Lys
437      115      120      125
440 Gly Met Tyr Tyr Leu Glu Glu His Gly Met Val His Arg Asn Leu Ala
441      130      135      140
444 Ala Arg Asn Val Leu Leu Lys Ser Pro Ser Gln Val Gln Val Ala Asp
445 145      150      155      160

```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/10/801,266

DATE: 08/19/2004
TIME: 09:54:39

Input Set : A:\PTO.LM.txt
Output Set: N:\CRF4\08192004\J801266.raw

Invalid Line Length:

The rules require that a line not exceed 72 characters in length. This includes spaces.

Seq#:1; Line(s) 7

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/801,266

DATE: 08/19/2004

TIME: 09:54:39

Input Set : A:\PTO.LM.txt

Output Set: N:\CRF4\08192004\J801266.raw